

**COPY**



**PCT OPERATIONS**

**FACSIMILE TRANSMISSION COVER SHEET**

DATE: 3/14/02

TO: Dianne Rees

TELEPHONE: (617) 951-3351

FAX NO.: (617) 439-4170

FROM: John Anderson

TELEPHONE: (703) 308-9116

FAX NO.: 703-308-4785 OR 703-305-3230

MESSAGE: Here is the SEQ Error Report.

NUMBER OF PAGES 10 (INCLUDING THIS PAGE)

5630

BIOTECHNOLOGY  
SYSTEMS  
BRANCH

## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/019,455  
Source: Pu/10  
Date Processed by STIC: 1/22/2002

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.  
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY.

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.  
PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)  
PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail. Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom, including:

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm>), **EFS Submission User Manual - ePAVE**
2. U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:  
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7<sup>th</sup> Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202  
Or  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

**Raw Sequence Listing Error Summary**

**ERROR DETECTED**      **SUGGESTED CORRECTION**      **SERIAL NUMBER:** 10/019,455

**ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE**

1  Wrapped Nucleic  
Wrapped Aminos  
The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2  Invalid Line Length  
The rules require that a line not exceed 72 characters in length. This includes white spaces.

3  Misaligned Amino  
Numbering  
The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4  Non-ASCII  
The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5  Variable Length  
Sequence(s) \_\_\_\_\_ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6  PatentIn 2.0  
"bug"  
A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) \_\_\_\_\_. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7  Skipped Sequences  
(OLD RULES)  
Sequence(s) \_\_\_\_\_ missing. If intentional, please insert the following lines for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
This sequence is intentionally skipped

8  Skipped Sequences  
(NEW RULES)  
Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.  
Sequence(s) 319 missing. If intentional, please insert the following lines for each skipped sequence.  
<210> sequence id number  
<400> sequence id number  
000

9  Use of n's or Xaa's  
(NEW RULES)  
Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10  Invalid <213>  
Response  
Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence  
11  Use of <220>  
Sequence(s) 1-2,7-8,13-22,27-28,32-38,42-45 missing the <220> "Feature" and associated numeric identifiers and responses.  
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

12  PatentIn 2.0  
"bug"  
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13  Misuse of n  
n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.



PCT10

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/019,455

DATE: 01/22/2002  
TIME: 10:38:55

Input Set : A:\56804 Sequence Listing.txt  
Output Set: N:\CRF3\01182002\J019455.raw

pp1-4  
Does Not Comply  
Corrected Diskette Needed

3 <110> APPLICANT: Takeda Chemical Industries, Ltd.  
4 <120> TITLE OF INVENTION: Novel Polypeptide and its Use  
5 <130> FILE REFERENCE: 2622WOOP  
6 <140> CURRENT APPLICATION NUMBER: US/10/019,455  
C--> 6 <141> CURRENT FILING DATE: 2001-10-28  
6 <150> PRIOR APPLICATION NUMBER: (1999-06-30) This goes next to L151  
W--> 7 <151> PRIOR FILING DATE: (JP 11-186718) This goes on L150 line  
WOK 8 <160> NUMBER OF SEQ ID: 48 50 (include  
skipped sequences)

Does Not Comply  
Corrected Diskette Needed

## ERRORED SEQUENCES

E--> 25 <210> SEQ ID NO: 4  
E--> 83 <210> SEQ ID NO: 10  
173 <210> SEQ ID NO: 19  
174 <211> LENGTH: 2019

seq. 3 missing - see p. 3  
seq. 9 missing - see p. 4

175 <212> TYPE: DNA  
176 <213> ORGANISM: Artificial Sequence  
W--> 177 <220> FEATURE:  
178 <223> OTHER INFORMATION:

W--> 179 <400> SEQUENCE: 19  
E--> 180 acccccctggc ccctctgga  
364 <210> SEQ ID NO: 36  
365 <211> LENGTH: 3936

366 <212> TYPE: DNA  
367 <213> ORGANISM: Artificial Sequence

W--> 368 <220> FEATURE:  
369 <223> OTHER INFORMATION:

W--> 370 <400> SEQUENCE: 36

E--> 371 cgaattccca ccatggtgtg gtcccccagtg ctcctt use lower-case letters 36  
426 <210> SEQ ID NO: 42

427 <211> LENGTH: 30

428 <212> TYPE: DNA

429 <213> ORGANISM: Artificial Sequence

W--> 430 <220> FEATURE:

431 <223> OTHER INFORMATION:

W--> 432 <400> SEQUENCE: 42

E--> 433 caccaggatg agatggaaat tgtgggttat use lower-case 30 use lower-case letters 30 insert cumulative base total at right margin of each line

434 <210> SEQ ID NO: 43

435 <211> LENGTH: 30

436 <212> TYPE: DNA

437 <213> ORGANISM: Artificial Sequence

W--> 438 <220> FEATURE:

Suggestion:  
Consult  
Sequence Rules

The types of errors shown exist throughout the Sequence Listing. Please check all sequences for similar errors.

20 19

see item 11  
on Error Summary  
sheet

item 11

item 11

(see p. 2)

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/019,455

DATE: 01/22/2002

TIME: 10:38:55

Input Set : A:\56804 Sequence Listing.txt  
Output Set: N:\CRF3\01182002\J019455.raw439 <223> OTHER INFORMATION: *item 11*

W--&gt; 440 &lt;400&gt; SEQUENCE: 43

E--&gt; 441 gggttatttc cccagcaact tggtttagaga

*30 ← insert*

442 &lt;210&gt; SEQ ID NO: 44

443 &lt;211&gt; LENGTH: 29

444 &lt;212&gt; TYPE: DNA

445 &lt;213&gt; ORGANISM: Artificial Sequence

W--&gt; 446 &lt;220&gt; FEATURE:

447 &lt;223&gt; OTHER INFORMATION:

W--&gt; 448 &lt;400&gt; SEQUENCE: 44

E--&gt; 449 agacacactc ctcatctgca cacaacttc

*item 11*

450 &lt;210&gt; SEQ ID NO: 45

451 &lt;211&gt; LENGTH: 30

452 &lt;212&gt; TYPE: DNA

453 &lt;213&gt; ORGANISM: Artificial Sequence

W--&gt; 454 &lt;220&gt; FEATURE:

455 &lt;223&gt; OTHER INFORMATION:

W--&gt; 456 &lt;400&gt; SEQUENCE: 45

E--&gt; 457 ctcctcatct gcacacacaact tcttagaaga

*29 ←*

521 &lt;210&gt; SEQ ID NO: 50

522 &lt;211&gt; LENGTH: 18

523 &lt;212&gt; TYPE: PRT

524 &lt;213&gt; ORGANISM: Rat

W--&gt; 525 &lt;400&gt; SEQUENCE: 50

526 Met Ala Arg Ile Leu Ile Leu Leu Gly Gly Leu Val Ala Leu Cys

527 5 10 15

528 Ala Gly

E--&gt; 529 10 delete - number the amino acids under every 5 amino acids

535 1/19

539 BOS2\_186291.1

*delete*

10/019,455 3

use lower-case letters for nucleotides when using  
new sequence  
rule format

<400> 2

CAGCGTGTGT ACCAGGAAGC TACCAA

<210> 4

26

where is  
sequence 3?

If intentionally  
skipped, use  
the format shown  
in item 8 on Env  
Summary Sheet

10/09/455 4

use lower-case letters

<400> 8

AACTTGGTGA AGGAGCAGCG TGTA

<210> 10

24

→ where is  
Sequence 9?

VERIFICATION SUMMARY  
PATENT APPLICATION: US/10/019,455

DATE: 01/22/2002  
TIME: 10:38:56

Input Set : A:\56804 Sequence Listing.txt  
Output Set: N:\CRF3\01182002\J019455.raw

L:4 M:283 W: Missing Blank Line separator, <120> field identifier  
L:5 M:283 W: Missing Blank Line separator, <130> field identifier  
L:6 M:270 C: Current Application Number differs, Replaced Current Application No  
L:6 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:7 M:256 W: Invalid Numeric Header Field, Wrong Prior FILING DATE:YYYY-MM-DD  
L:8 M:283 W: Missing Blank Line separator, <160> field identifier  
L:9 M:283 W: Missing Blank Line separator, <210> field identifier  
L:13 M:283 W: Missing Blank Line separator, <220> field identifier  
L:15 M:283 W: Missing Blank Line separator, <400> field identifier  
L:16 M:112 C: (48) String data converted to lower case,  
L:21 M:283 W: Missing Blank Line separator, <220> field identifier  
L:23 M:283 W: Missing Blank Line separator, <400> field identifier  
L:24 M:112 C: (48) String data converted to lower case,  
L:25 M:214 E: (33) Seq.# missing, SEQ ID NO:3  
L:29 M:283 W: Missing Blank Line separator, <400> field identifier  
L:30 M:112 C: (48) String data converted to lower case,  
M:112 Repeated in SeqNo=4  
L:41 M:283 W: Missing Blank Line separator, <400> field identifier  
L:50 M:283 W: Missing Blank Line separator, <400> field identifier  
L:71 M:283 W: Missing Blank Line separator, <220> field identifier  
L:73 M:283 W: Missing Blank Line separator, <400> field identifier  
L:74 M:112 C: (48) String data converted to lower case,  
L:79 M:283 W: Missing Blank Line separator, <220> field identifier  
L:81 M:283 W: Missing Blank Line separator, <400> field identifier  
L:82 M:112 C: (48) String data converted to lower case,  
L:83 M:214 E: (33) Seq.# missing, SEQ ID NO:9  
L:87 M:283 W: Missing Blank Line separator, <400> field identifier  
L:88 M:112 C: (48) String data converted to lower case,  
M:112 Repeated in SeqNo=10  
L:99 M:283 W: Missing Blank Line separator, <400> field identifier  
L:108 M:283 W: Missing Blank Line separator, <400> field identifier  
L:129 M:283 W: Missing Blank Line separator, <220> field identifier  
L:131 M:283 W: Missing Blank Line separator, <400> field identifier  
L:132 M:112 C: (48) String data converted to lower case,  
L:137 M:283 W: Missing Blank Line separator, <220> field identifier  
L:139 M:283 W: Missing Blank Line separator, <400> field identifier  
L:140 M:112 C: (48) String data converted to lower case,  
L:145 M:283 W: Missing Blank Line separator, <220> field identifier  
L:147 M:283 W: Missing Blank Line separator, <400> field identifier  
L:148 M:112 C: (48) String data converted to lower case,  
L:153 M:283 W: Missing Blank Line separator, <220> field identifier  
L:155 M:283 W: Missing Blank Line separator, <400> field identifier  
L:156 M:112 C: (48) String data converted to lower case,  
L:161 M:283 W: Missing Blank Line separator, <220> field identifier  
L:163 M:283 W: Missing Blank Line separator, <400> field identifier  
L:164 M:112 C: (48) String data converted to lower case,  
L:169 M:283 W: Missing Blank Line separator, <220> field identifier  
L:171 M:283 W: Missing Blank Line separator, <400> field identifier

VERIFICATION SUMMARY  
PATENT APPLICATION: US/10/019,455

DATE: 01/22/2002  
TIME: 10:38:56

Input Set : A:\56804 Sequence Listing.txt  
Output Set: N:\CRF3\01182002\J019455.raw

L:172 M:112 C: (48) String data converted to lower case,  
L:177 M:283 W: Missing Blank Line separator, <220> field identifier  
L:179 M:283 W: Missing Blank Line separator, <400> field identifier  
L:180 M:254 E: No. of Bases conflict, LENGTH:Input:20 Counted:19 SEQ:19  
M:112 Repeated in SeqNo=19  
L:180 M:252 E: No. of Seq. differs, <211>LENGTH:Input:20 Found:19 SEQ:19  
L:185 M:283 W: Missing Blank Line separator, <220> field identifier  
L:187 M:283 W: Missing Blank Line separator, <400> field identifier  
L:188 M:112 C: (48) String data converted to lower case,  
L:193 M:283 W: Missing Blank Line separator, <220> field identifier  
L:195 M:283 W: Missing Blank Line separator, <400> field identifier  
L:196 M:112 C: (48) String data converted to lower case,  
L:201 M:283 W: Missing Blank Line separator, <220> field identifier  
L:203 M:283 W: Missing Blank Line separator, <400> field identifier  
L:204 M:112 C: (48) String data converted to lower case,  
L:209 M:283 W: Missing Blank Line separator, <400> field identifier  
L:210 M:112 C: (48) String data converted to lower case,  
M:112 Repeated in SeqNo=23  
L:220 M:283 W: Missing Blank Line separator, <400> field identifier  
L:239 M:283 W: Missing Blank Line separator, <400> field identifier  
L:240 M:112 C: (48) String data converted to lower case,  
M:112 Repeated in SeqNo=25  
L:250 M:283 W: Missing Blank Line separator, <400> field identifier  
L:269 M:283 W: Missing Blank Line separator, <220> field identifier  
L:271 M:283 W: Missing Blank Line separator, <400> field identifier  
L:272 M:112 C: (48) String data converted to lower case,  
L:277 M:283 W: Missing Blank Line separator, <220> field identifier  
L:279 M:283 W: Missing Blank Line separator, <400> field identifier  
L:280 M:112 C: (48) String data converted to lower case,  
L:285 M:283 W: Missing Blank Line separator, <400> field identifier  
L:286 M:112 C: (48) String data converted to lower case,  
M:112 Repeated in SeqNo=29  
L:306 M:283 W: Missing Blank Line separator, <400> field identifier  
L:307 M:112 C: (48) String data converted to lower case,  
M:112 Repeated in SeqNo=30  
L:327 M:283 W: Missing Blank Line separator, <400> field identifier  
L:336 M:283 W: Missing Blank Line separator, <220> field identifier  
L:339 M:112 C: (48) String data converted to lower case,  
L:347 M:112 C: (48) String data converted to lower case,  
L:355 M:112 C: (48) String data converted to lower case,  
L:363 M:112 C: (48) String data converted to lower case,  
L:371 M:112 C: (48) String data converted to lower case,  
L:371 M:252 E: No. of Seq. differs, <211>LENGTH:Input:39 Found:36 SEQ:36  
L:379 M:112 C: (48) String data converted to lower case,  
L:387 M:112 C: (48) String data converted to lower case,  
L:410 M:112 C: (48) String data converted to lower case,  
M:112 Repeated in SeqNo=40  
L:420 M:112 C: (48) String data converted to lower case,  
M:112 Repeated in SeqNo=41

## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/019,455

DATE: 01/22/2002

TIME: 10:38:56

Input Set : A:\56804 Sequence Listing.txt  
Output Set: N:\CRF3\01182002\J019455.raw

L:433 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:30 SEQ:42  
M:112 Repeated in SeqNo=42  
L:441 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:30 SEQ:43  
M:112 Repeated in SeqNo=43  
L:449 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:29 SEQ:44  
M:112 Repeated in SeqNo=44  
L:457 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:30 SEQ:45  
M:112 Repeated in SeqNo=45  
L:463 M:112 C: (48) String data converted to lower case,  
M:112 Repeated in SeqNo=46  
L:496 M:112 C: (48) String data converted to lower case,  
M:112 Repeated in SeqNo=48  
L:529 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:50